

BERICHTIGTE FASSUNG

Serial No. 10/579,248
Group No. 1752
Confirmation No.7812

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
23. Mai 2002 (23.05.2002)

PCT

(10) Internationale Veröffentlichungsnummer
WO 02/040511 A2

(51) Internationale Patentklassifikation⁷: C07K 14/00

(21) Internationales Aktenzeichen: PCT/EP01/12937

(22) Internationales Anmeldedatum:
8. November 2001 (08.11.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
100 55 789.9 10. November 2000 (10.11.2000) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): PHARMA-ZENTRALE GMBH [DE/DE];
Loerfeldstrasse 20, 58313 Herdecke (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): HACKER, Jörg
[DE/DE]; Edith-Stein-Strasse 6, 97218 Gerbrunn (DE).
BLUM-OEHLER, Gabriele [DE/DE]; Goethestrasse 3,
97072 Würzburg (DE). JUNG, Günther [DE/DE]; Ob
der Grafenhalde 5, 72076 Tübingen (DE). HANTKE,
Klaus [DE/DE]; Beckmannweg 7/1, 72076 Tübingen
(DE). PATZER, Silke [DE/DE]; Mörikestrasse 16, 71134
Aidlingen (DE). MORENO, Felipe [ES/ES]; C/ Mar-
bella N°.16, 10B, E-28034 Madrid (ES). BAQUERO,
Fernando [ES/ES]; C/ Ardemans N°. 62, 30A, E-28028
Madrid (ES). BAQUERO, Rosario [ES/ES]; C/ La Baneza
N°. 43 Portal 2, 90A, E-28035 Madrid (ES). BRAVO,
Daniel [ES/ES]; C/Madero N°. 11, 10A, E-28412 Cerceda
(ES). SONNENBORN, Ulrich [DE/DE]; Eichenweg
6, 44799 Bochum (DE). SCHULZE, Jürgen [DE/DE];
Alice-Bloch-Strasse 7, 14558 Bergholz-Rehbrücke (DE).

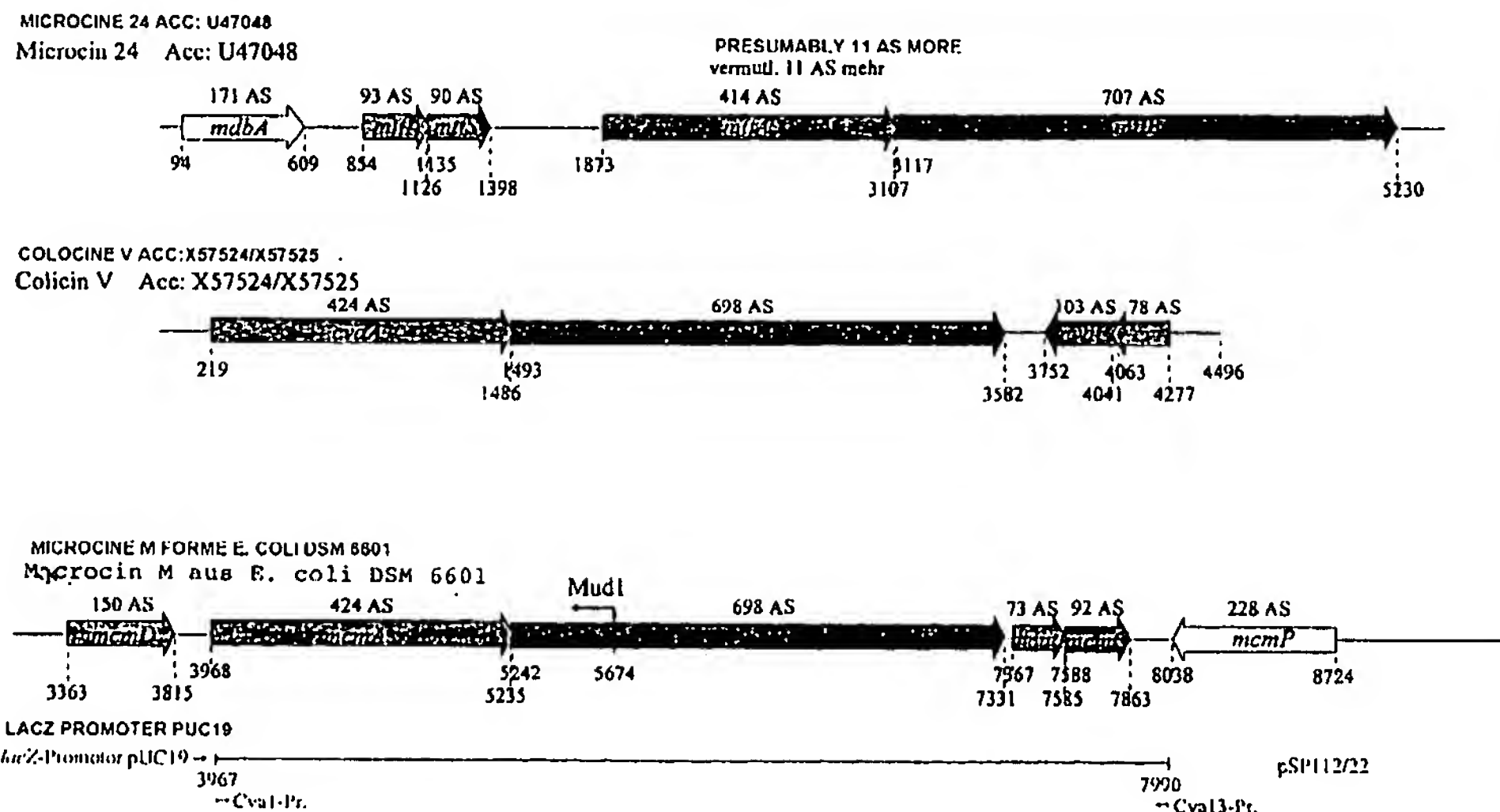
[Fortsetzung auf der nächsten Seite]

(54) Title: DNA SEQUENCE AND MICROCIN FROM ESCHERICHIA COLI STRAIN DSM 6601

(54) Bezeichnung: DNA-SEQUENZ UND MICROCIN AUS ESCHERICHIA COLI STAMM DSM 6601

ORGANISATION OF THE GENE CLUSTERS OF THE MCM REGION FROM E. COLI DSM 6601
WITH THE GENE SECTIONS NECESSARY FOR THE SYNTHESIS PROCESSING AND EXPRESSION
OF MICROCIN (BELOW) AND COMPARISON WITH THE GENE CLUSTERS FOR THE PRODUCTION
OF MICROCIN 24 (ABOVE) AND COLICIN V (MIDDLE)

Organisation des Gen-Clusters der mcm-Region aus E. coli DSM 6601 mit den
für die Microcin-Synthese, -Prozessierung und -Ausscheidung notwendigen
Genabschnitten (unten) und Vergleich mit den Gen-Clustern für die Produk-
tion von Microcin 24 (oben) und Colicin V (mitte)



(57) Abstract: The invention relates to a DNA sequence with the nucleotide sequence given in figure 1 and the peptide encoded thereby, derivatives of said peptide, and use of the DNA sequence, the peptide and the peptide derivative.

(57) Zusammenfassung: Die Erfindung betrifft eine DNA-Sequenz mit der in Abbildung 1 dargestellten Nucleotidfolge und das dadurch codierte Peptid, Derivate des Peptids sowie die Verwendung der DNA-Sequenz, des Peptids sowie der Peptidderivate.



PROPPERT, Hans [DE/DE]; Rosenstrasse 102, 58095 Hagen (DE).

ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR),
OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

(74) **Anwalt: SIEWERS, Gescha**; Dres. Harmsen & Utescher
PP., Alter Wall 55, 20457 Hamburg (DE).

Veröffentlicht:

— *ohne internationalen Recherchenbericht und erneut zu
veröffentlichen nach Erhalt des Berichts*

(81) **Bestimmungsstaaten (national):** AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,
CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZW.

(48) **Datum der Veröffentlichung dieser berichtigten**

Fassung: 25. Juli 2002

(15) **Informationen zur Berichtigung:**

siehe PCT Gazette Nr. 30/2002 vom 25. Juli 2002, Section
II

(84) **Bestimmungsstaaten (regional):** ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches Patent (AT, BE, CH, CY, DE, DK,

*Zur Erklärung der Zweibuchstaben-Codes und der anderen
Abkürzungen wird auf die Erklärungen ("Guidance Notes on
Codes and Abbreviations") am Anfang jeder regulären Ausgabe
der PCT-Gazette verwiesen.*

DNA-Sequenz und Microcin aus *Escherichia coli* Stamm DSM 6601

Die Erfindung betrifft eine DNA-Sequenz und ein neues hierdurch codiertes Microcin aus *Escherichia coli* Stamm DSM 6601.

5

Escherichia coli ist ein gramnegatives Bakterium, das in der menschlichen und tierischen Darmflora, aber auch extraintestinal vorkommt. *E. coli* tritt in zahlreichen Varianten auf, die sich hinsichtlich der Kapsel-, Oberflächen- und Flagellenantigene unterscheiden und in zahlreiche serologische
10 Typen unterteilt werden können. Die Einordnung nach den Serotypen läßt allerdings keine Aussage über die unterschiedliche Virulenz der Keime zu. Vertreter ein- und desselben Serotyps können sowohl im menschlichen als auch im tierischen Körper ein unterschiedliches Pathogenitätspotential besitzen, das im Extremfall von avirulent bis hochgradig pathogen reichen
15 kann. Der *E. coli* Stamm DSM 6601 gehört zu der Serogruppe O6:K5 und wird als nicht human und nicht tierpathogen bewertet.

Als Bakteriocine bezeichnet man spezifische Proteine, die von bestimmten Bakterien produziert werden und mit abtötender Wirkung gegen andere
20 Stämme derselben oder einer nahe verwandten Bakterienart gerichtet sind. Es sind bereits über zwanzig Colicine aus *E. coli* bekannt, die jeweils hohe antibakterielle Spezifität aufweisen. Zu den Bakteriocinen zählen auch die sogenannten Microcine, die u.a. auch von bestimmten *E. coli*-Stämmen produziert werden. Bakteriocine werden, abgesehen von ihrer
25 antibiotischen Wirkung, für die Taxonomie von Bakterienstämmen eingesetzt.

Es war bekannt, daß *E. coli* Stamm DSM 6601 antibiotische Aktivität aufweist, die dem Vorhandensein eines Microcins zugeschrieben wurde,

das im folgenden als Microcin M (McmC) bezeichnet wird. Versuche, das Peptid durch Anreicherung aus Flüssigkulturen oder durch Extraktion von Kulturen auf Agarplatten oder aus dem Agar anzureichern, brachten keine verwertbaren Ergebnisse. Es wurde daher versucht, das Microcin mit Hilfe

5 genetischer Methoden zu identifizieren. Mit dem oben genannten Stamm von *E. coli* wurde eine Mu-Mutagenese durchgeführt. Der Phage Mu kann sich nach Infektion einer Bakterienzelle wie ein Transposon verhalten und an einem beliebigen Ort in der DNA integrieren und dadurch eine Mutation hervorrufen. Im vorliegenden Fall wurde der Phage Mud 1 nach der

10 Vorschrift von Casadaban (Casadaban und Cohen, 1979, Proc. Natl. Acad. Sci. 76, 4530-4533) verwendet. Durch die Mutagenese wurde eine Mutante isoliert, die kein Microcin mehr produzierte. Aus der chromosomalen DNA dieser Mutante wurde das für die Microcinsynthese kodierende DNA-Fragment isoliert und nach Insertion in den Vektor

15 pUC18 und PCR-Amplifikation kloniert und sequenziert.

Überraschenderweise wurde anhand der Sequenzanalyse festgestellt, daß die *mcm*-Region aus *E. coli* DSM 6601 einen dem Colicin V ähnlichen Hemmstoff kodiert, der im folgenden Microcin M (McmC) genannt wird.

20 Die Abbildung 1 zeigt die entsprechende Sequenz der *mcm*-Region.

Die Abbildung 1a zeigt in der Übersicht die in diesem Bereich aufgrund der DNA-Sequenz vorhergesagten Gene. Die Genprodukte McmA und McmB zeigen hohe Sequenzähnlichkeit zu den Exportproteinen für Colicin

25 V und Microcin 24. Es folgt das Gen *mcml*, dessen Produkt nur geringe Ähnlichkeit zu den Immunitätsproteinen der beiden obigen Microcine zeigt. Direkt im Anschluß wird das Microcin M (McmC) kodiert. Das darauf folgende Gen *mcmP* wird in entgegengesetzter Richtung abgelesen und wirkt aufgrund seiner Homologien zum *cvpA*-Gen (Colicin V Produktion)

bei der Expression von MmcC mit. Stromauf vom *mcmA*-Gen liegt das Gen *mcmD*, dessen Produkt Ähnlichkeit mit den fettsäureübertragenden Proteinen (Transacylasen) aufweist, die auch in die posttranslationale Modifikation der *E. coli*-Hämolyse involviert sind.

5

Zur Erläuterung wird darauf hingewiesen, daß das Protein CvaB zur Familie der MDR-Proteine (Multi-Drug-Resistance) gehört, die ATP-abhängig den Export oder Import von kleinen Substanzen, aber auch den Export von Proteinen katalysieren. Es handelt sich um Proteine mit 6 oder
 10 8 hydrophoben Helices in der Membran und einer Domäne mit einer ATP-spaltenden Aktivität, womit der Transportvorgang energetisiert wird. Die ATP-spaltende Aktivität kann auch als separates Protein vorliegen. CvaB und zehn weitere verwandte Proteine besitzen im Gegensatz zu anderen MDR-Proteinen eine N-terminale Domäne mit Cysteinproteinase-Aktivität
 15 im Cytoplasma. Diese Proteinase spaltet beim Export eine kleine Leaderpeptidsequenz vom N-terminalen Ende des Colicins V ab und zwar nach der Gly-Gly Erkennungssequenz (siehe Pfeil bei dem Homologieversuch in Abbildung 1a). Eine solche Gly-Gly Sequenz findet sich auch in der abgeleiteten Sequenz für das unreife MmcC, einem
 20 Peptid aus 92 Aminosäuren. Für das reife MmcC-Peptid ist somit eine Kettenlänge von 77 Aminosäuren vorgegeben.

Das Protein CvaA wird über eine alpha-Helix in der Cytoplasmamembran verankert und dient als Verbindungsprotein (Connector) zur äußeren
 25 Membran. In der äußeren Membran bindet CvaA an das TolC Protein. Soweit bekannt, bilden die Proteine CvaB, CvaA und TolC einen Proteinkomplex, der den Export von Colicin V über die Cytoplasmamembran und die äußere Membran in einem Schritt katalysiert. Sehr ähnliche Exportapparate hat man z.B. für das *E. coli*-

Hämolysin oder für Proteinasen aus *Erwinia*-Arten gefunden. Aufgrund der hohen Ähnlichkeit zu CvaA und CvaB dürften das McM A und McM B dieselben Funktionen beim Export und bei der Reifung von McM C erfüllen.

- 5 In den nachfolgenden Tabellen ist die prozentuale Identität der Microcine, der Exportproteine und der Immunitätsproteine sowie die prozentuale Identität im Vergleich zu anderen verwandten Proteinen zusammengestellt:

10 Tabelle 1 a: Identität der Microcine im Einzelvergleich in Prozent

	ColV	Mc mC	Mic24	MicE492
ColV	100	35	19	15
Mc mC		100	10	28
Mic24			100	45
MicE492				100

Tabelle 1 b: Identität der dem CvaA ähnlichen Exportproteine im Einzelvergleich in Prozent

15

	CvaA	Mc mA	MtfA
CvaA	100	97	71
Mc mA		100	71
MtfA			100

Tabelle 1c: Identität der dem CvaB verwandten Exportproteine in Prozent

	CvaB	Mc mB	MtfB
CvaB	100	89	73
Mc mB		100	75

MtfB			100
------	--	--	-----

Tabelle 1 d: Prozentuale Identität im Vergleich zu anderen McmD verwandten Proteinen. Abkürzungen: HlyCneim ein HlyC ähnliches Protein aus Neisseria meningitidis; HlyCec HlyC Protein aus Escherichia coli

	McmD	HlyCneim	HlyCec
McmD	100	25	12
HlyCneim		100	11
HlyCec			100

Tabelle 1 e: Prozentuale Identität der Immunitätsproteine im Einzelvergleich. Abkürzungen: Cval – Immunitätsprotein von ColV, Mcml – Immunitätsprotein von Microcin McM, Mtfl – Immunitätsprotein von Microcin M24

	Cval	Mcml	Mtfl
Cval	100	23	14
Mcml		100	22
Mtfl			100

Aufgrund der DNA-Analyse ergibt sich für die neue Substanz ein Peptid mit 77 Aminosäuren, das aufgrund seiner Größe den Microcinen zugerechnet wird. Dieses reife Microcin (McmC) zeigt aber nur geringe Sequenzhomologie zu bekannten Polypeptiden wie z.B. Colicin V.

Zur Synthese und zur näheren Untersuchung des Peptids wurden überlappende Oligopeptide mit jeweils fünfzehn Aminosäuren synthetisiert, also 1 bis 15, 5 bis 20, 10 bis 25 usw., und zwar durch Oligopeptidsynthese unter Anwendung der Fmoc-Strategie auf einem Polystyrol-1% Divinylbenzolharz mit 2-Chlortriphenylmethylchlorid als Linker.

Die Gesamtsynthese des Microcin-Peptids kann in an sich bekannter Weise durch Segmentkondensation auf dem Harz unter Verwendung von drei bis vier Peptid-Segmenten und Einsatz des ABI-Peptidsynthesizers erfolgen, wobei vorzugsweise so vorgegangen wird, daß mit dem an das Harz gebundenen Segment I., bestehend aus den Aminosäuren 77 bis 54, die gereinigten Segmente II. (Aminosäuren 53 bis 24) und III. (Aminosäuren 23 bis 1) gekoppelt werden.

Die Herstellung von Lipopeptidmischungen zur Gewinnung von Antiseren erfolgt in an sich bekannter Weise, wobei Pam₃Cys-Lys-Ser-Peptide zur Immunisierung verwendet werden können.

5

Zur Beschichtung von Streptavidinplatten für ELISA-Teste werden in an sich bekannter Weise Biotinyl-Aca-Aca-Peptide verwendet, wobei diese Beschichtungen sowohl mit den entsprechenden Einzelpeptiden als auch mit einer Mischung von Peptiden durchgeführt werden können.

10

Im folgenden wird die Erfindung anhand der Beispiele näher erläutert:

Beispiel 1

15 Identifizierung und Sequenzanalyse der mcm-Region aus E. coli DSM 6601

Nach dem Verfahren von Casadaban et al. wurde eine Mutante, die kein Microcin mehr produzierte, im Plattentest mit dem Indikatorstamm H1941
20 identifiziert. Chromosomale DNA dieser Mutante wurde isoliert und mit dem Restriktionsenzym HincII verdaut, danach mit dem mittels SmaI geschnittenen Vektor pUC18 ligiert und mit den Primern MuD1 (spezifisch für ein Mu-DNA-Ende) und UNI (spezifisch für den ligierten pUC18 Vektor) in einer PCR amplifiziert. Das DNA-Fragment wurde nach bekannten
25 Verfahren (J. Sambrook, E.F. Fritsch, T. Maniatis; Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989) kloniert und sequenziert. Es stellte sich heraus, daß dieses DNA-Fragment hohe Sequenzähnlichkeit zu dem Colicin-V-Exportprotein B aufwies. Klonierung und Sequenzanalyse des gesamten Bereiches zeigten, daß diese Region
30 einen dem Colicin V ähnlichen Hemmstoff kodiert.

Beispiel 2

Synthese des Microcins

35

Die Aminosäuresequenz des Microcin MmcC läßt sich aus der DNA der Region, die dieses Peptid kodiert, ableiten und ergibt sich aus der beigefügten Abbildung 3.

40

Zur Synthese des Peptids werden in an sich bekannter Weise (G. Jung, Ed.; Combinatorial Peptide and Non-Peptide Libraries, Wiley VCH, Weinheim, 1996) Oligopeptide mit fünfzehn Aminosäuren synthetisiert, die überlappend die Aminosäuren 1 bis 15, 5 bis 20, 10 bis 25, 15 bis 30 usw. aufweisen. Die Synthese erfolgt unter Verwendung von Fmoc-Schutzgruppen auf einem Polystyrol-1%-Divinylbenzolharz nach dem Merryfield-Verfahren mit 2-Chlortriphenylmethylchlorid als Linker.

Nach hinreichend langer Synthesedauer werden in den jeweiligen Ansätzen die Peptide vom Harz gelöst und unter Verwendung von HPLC und ES-MS-Spektroskopie untersucht. Die Sekundärstruktur kann zum Teil durch Untersuchung des Circular dichroismus festgestellt werden.

Die Gesamtsynthese des Peptids aus siebenundsiebzig Aminosäuren erfolgt unter Verwendung eines ABI-Peptidsynthesizers unter Kondensation von Peptidsegmenten auf dem Harz, wobei vorzugsweise mit dem harzgebundenen Peptidsegment I. (Aminosäuren 77 bis 54) angefangen wird und dann die gereinigten Segmente II. (Aminosäuren 53 bis 24) und III. (Aminosäuren 23 bis 1) gekoppelt werden.

Die freien Peptide können über HPLC gereinigt werden, wobei die Reinheit für die Derivatisierung etwa 90% betragen sollte.

Für die Derivatisierung zur Herstellung von Lipopeptiden wird nach publizierten Verfahren vorgegangen (K.-H. Wiesmüller, G. Jung, G. Hess; Vaccine 7: 29 – 33, 1989; G. Jung, K.-H. Wiesmüller, G. Becker, H.-J. Bühring, WG Bessler; Angewandte Chemie [Internat. Ed.] 24: 872 – 873, 1985). Die Herstellung von für ELISA-Teste geeigneten biotinylierten Peptiden erfolgt ebenfalls nach bereits publizierten Verfahren (G. Jung, Ed.; Combinatorial Peptide and Non-Peptide Libraries, Wiley VCH, Weinheim, 1996).

Patentansprüche

1. DNA-Sequenz mit der in Abbildung 1 dargestellten Nucleotidfolge.
- 5 2. Verwendung der DNA-Sequenz nach Anspruch 1 in der mikrobiologischen Analytik und/oder Diagnostik.
3. Verwendung der DNA-Sequenz nach Anspruch 1 in der Biotechnik als Expressionsvektor oder als Bestandteil von Multi-Copy-Plasmiden zur Herstellung des antimikrobiellen Wirkstoffes.
- 10 4. Peptid mit der in Abbildung 2 dargestellten Aminosäuresequenz.
5. Verwenung des Peptids nach Anspruch 4 zur Herstellung von Antipeptid – Antikörpern.
- 15 6. Verwendung des Peptides nach Anspruch 4 als antimikrobieller Wirkstoff zu mikrobiologischen oder medizinischen Zwecken.
- 20 7. Oligopeptide mit den in Abbildung 3 dargestellten Aminosäuresequenzen.
8. Verwendung der Oligopeptide nach Anspruch 7 zur Herstellung von Antipeptid-Antikörpern.
- 25 9. Verwendung der Oligopeptide nach Anspruch 7 als antimikrobielle Wirkstoffe zu mikrobiologischen oder medizinischen Zwecken.

10. Lipopeptide nach Anspruch 4 oder 7, gekennzeichnet durch Kopplung zu Pam₃Cys-Lys-Ser-Peptiden.
11. Peptide nach Anspruch 4 oder 7, gekennzeichnet durch Kopplung zu Biotinyl-Aca-Aca-Peptiden.
12. Verwendung der Lipopeptide nach Anspruch 10 zur Herstellung von Antiseren.
13. Verwendung der Peptide nach Anspruch 11 zur Herstellung von ELISA-Tests.

Abbildung 1 Die 5' Basenpaare lange DNA-Sequenz der kompletten mcr Region aus E. coli DSM 6601

ID MCM PRELIMINARY; DNA; 5381 BP.
 SQ SEQUENCE 5381 BP; 1601 A; 971 C; 1269 G; 1540 T; 0 OTHER;

```

ATAAGGGAAA TAGTAATCAT GTCTTATATA AGGGAAACCA TCAGAGGAAA AGATGAATGG
ACTGTTTATG AACAGATCGG TTTTGC GGTC AGTTGTATGC TCTACAATCG TAATTACAGT
CTGTATCCGG TGTTAACCAT TCAATACTGG ACTGAATATG CGATACAGCA TAATCAGATT
AAATTCCCTGT TTGATTCACG AGGTTTTCCA CTGGCGTATA TTACCTGGGC ATATCTTGAG
GCTGATACAG AAGCACGCCT GCTCAGGGAT CCAGAATTCA GGTTCATCC GTCTGAATGG
AATGAAGATG GAAGGATCTG GATTCTGGAT TTCTGTTGTA AACCAGGCTT TGGTCGAAAA
GTTATTGACT ATCTCATACA GCTTCAGCCA TGGGGAGAAG GAGAAGTACG ATGGTTAAGC
AGGCGAAAGA AAATTGTGAC ATACATCCCT GAGCGGCTGC ATAAAACGTA GTACCTCTGA
AGATACAGAG ATAATTGTAA ATTACGGGGT AAATGCATCG CTGATACTAT TTTGACAGGA
CTCTGTATTT TCTGGCTTTG ATTATAGCAT TTCGGTTAAC CAGTGAAGGT ATTACAGGAA
AAGTGTGAGT AAGAGGAGCT ATTTTGTTC GTCAGGATGC TTTAGAAAAAC AGAAAAATGA
AGTGGCAGGG ACGGGCAATA TTACTTCCCG GAATACCACT ATGGTTAATC ATGCTGGGAA
GCATAGTGTT TATTACGGCA TTTCTGATGT TCATTATTGT TGGTACCTAT AGCCGCCGTG
TTAATGTCAG TGGTGAGGTC ACAACCTGGC CAAGAGCTGT CAATATATAT TCAGGTGTAC
AGGGATTTGT TGTCAGGCAG TTTGTTTCATG AAGGACAGTT GATAAAAAA GGGGATCCTG
TTTATCTGAT TGACATCAGT AAAAGTACAC GTAGTGGTAT TGTCACCTGAT AATCATCGTC
GGGATATAGA AAACCAGCTG GTTCGTGTGG ACAAGATTAT TTCCCGTCTG GAAGAAAGTA
AAAAAATAAC GCTGGATACC CTGGA AAAAC AACGCCTGCA ATACACAGAT GCGTTTCTGC
GCTCATCAGA TATTATACAG CGTGCAGAGG AAGGGATAAA AATAATGAAA AACACATGG
AGAATTACAG AAACATCAG ACAAAAGGGC TGATTAATAA AGATCAGTTA ACTAACCAGG
TGGCATTATA TTATCAGCAA CAAAACAATC TTCTTAGCCT GAGCGGACAG AACGAACAGA
ATGCCCTGCA GATAACCACT CTGGAGAGTC AGATTACAGC TCAGGCAGCA GATTTTGATA
ACCGTATCTA CCAGATGGAA CTGCAACGGT ACGAGTTACA GAAAGAACTG GTTAACACTG
ATGTGGAGGG CGAAATCATT ATCCGGGCGT TGACTGACGG GAAAGTTGAC TCCCTGAGTG
TCACTGTCCG GCAAATGGTC AATCCCGGAG ACAACCTCT GCAGGTATT CCTGAGACA
TTGAAAATTA TTATCTTATT CTCTGGGTCC CGAATGATGC TGTTCCCTAT ATTTCCGGCTG
GTGACAAAGT GAATATTCGT TATGAAGCCT TTCCGGCAGA AAAATTTGGG CAGTTTTCTG
CTACGGTTAA AACTATATCC AGGACTCCTG CATCAACACA GGAAATGTTG ACCTATAAGG
GAGCACCACA GAATACGCCG GCGCCTCTG TTCCCTGGTA TAAAGTCATT GCGATGCCTG
AAAAGCAGAT TATCAGATAT GACGAAAAA ACCTCCCTCT GGAAAATGGA ATGAAAGCCG
AAAGTACACT ATTTCTGGAA AAAAGGCGTA TTTACCAGTG GATGCTTTCT CCTTTCTATG
ACATGAAACA CAGTGCAACA GGACCGCTCA ATGACTAACG GGAATTTTCA ACAAATTATA
AATCAGCTTG ATATGCGCTG GCGACGTCGT GTTCCGGTTA TTCATCAGAC GGAGACCGCT
GAATGTGGAC TGGCCTGCCT GGCAATGATA TGCGGTCATT TTGGTAAGAA TATTGACCTG
ATATCTCTTC GTCGTAAGTT TAATCTCTCG GCCCGTGGAG CAAACCTTGC AGGAATCAAT
GGAATAGCGG AGCAGCTGGG GATGATCACC CGGGCTCTTT CACTGGAGCT GGATGAACTT
GGTGCCCTCA AAATGCCGTG TATTCTCCAC TGGGATTTCA GTCACCTCGT CGTGCTGGTC
AGCGTAAAGC GTAAACGTTA TGTACTGCAT GATCCGGCCA GAGGAAGAAG ATATCTCGGT
CGGGAGGAAA TGAGCCGGTA TTTTACGGGC ATTGCACTTG AGGTCTGGCC CGGAGGTGAA
TTCCGGACGG AAACCCAGCA AACCCGCATA AGTCTCCGTT CACTGATTAA CAGTATTTAC
GGTATTAAAA GAACACTGGC GAAAATTTTC TGTCTGTCAG CTGTAATTGA AGCAATCAAT
CTGGTAATGC CGGTGGGAAC TCAGCTGGTT ATGGATCATG CGATTCCGGC GGGGGACCGA
GGGCTGCTGA CGCTTATTTT TGCTGGCCTG ATGTTCTTTA TATTGCTCAG GGCCGCGGTG
AGTATGCTGC GTGCATGGTC CTCACTGGTT ATGAGCACGC TCATCAATAT ACAGTGGCAG
TCGGGTCTGT TTAACCACTT TCTCAGACTG CCGCTGGCAT TTTTGAACG CCGTAAATTA
GGTGATATCC AGTCGCGTTT TGGCTCCCTT GACACTTTGA GGGCCACCTT TACCACCTGT
GTGGTTGGGG CAATCATGGA CAGTATTATG GTTGTGGGTG TTTTGTGAT GATGCTGTTA
TATGGAGGAT ATCTTACCTG GATAGTTCTC GGTTTTACCA TGGTTTACGT TCTTATTCGT
CTGGTGACAT ACGGCTATTA CCGGCAATA TCGGAAGAAA CTCTTGTCAG GGGGGCCCGG
GCCAGCTCCT ATTTTATGGA AAGCCTGTAT GGTATTGCCA CCGTAAAAAT CCAGGGTATG
GTCGGGATCC GGGGAACACA CTGGCTTAAC CTGAAAATAG ATGCGATCAA TTCAGGTATT
AAGTTAAACA GGATGGATT GCTCTTCGGG GGAATAAATA CTTTGTGTTG CGCCTGTGAT
CAGGTGGCGA TTTTATGGCT GGGTGCAAGC CTTGTGATCG ATAATCAGAT GACAATAGGG
ATGTTTGTGG CATTTGGTTC TTTTCGTGGG CAGTTTTCGG ATCGGTTTGC TTCGCTGACC
AGTTTTCTTC TTCAACTGAG AATAATGAGT CTGCATAATG AGCGCATTGC AGATATTGCA
CTACATGAAA AGGAAGAAA GAAACCGGAA ATTGAAATCG TTGCTGACAT GAGCCCGGTT
TCACTGGAAA CCACTGATTT AAGCTACCGG TATGACAGCC AGTCAGCACA GGTATTCAGT
GGTCTGAATT TGTCTGTGGC TCCGGGAGAA AGTGTGGCTA TAACTGGTGC CTCCGGTGCC
GGAAAAACCA CATTAATGAA AGTATTATGT GGACTGTTTG AACCAGATAG TGGAAAAGTA
CTGGTTAATG GCACGGATAT ACGTCAACTT GGAATAAATA ATTATCACCG TATGATAGCC
TGTGTTATGC AGGACGACCG GCTATTTTCA GGATCAATTC GTGAAAATAT CTGTGGGTTT
GCAGAAAGAA CAGACGACGA ATGGATGACA GAATGTGCCA GAGCAAGTCA TATTCATGAT
GTGATAATGA AAATGCCAAT GGGGTATGAA ACGTTAATAG GTGAACTGGG GGAAGGTCTT
TCCGGCGGTC AAAACAGCG TATATTCATT GCCCGAGCTT TATACCGGAA ACCTGGAATA
TTATTTATGG ATGAGGCTAC AAGTCTCTT GATACAGAAA GTGAACGTTT CGTTAATGCT

```

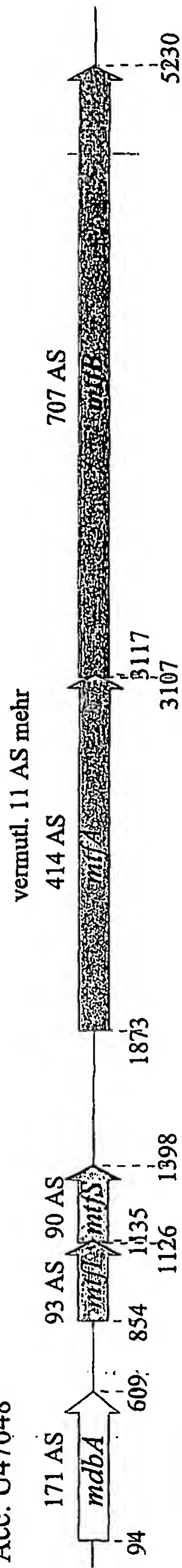
```

GCCATAAAAA AAATGAATAT CACCCGGGTG ATTATTGCAC ACAGAGAAAC TACGTTGAGA
ACTGTTGACA GGATTATTTT TATTTAAAAAT CCACTGGTGT AACTTTGTAA GGAGTTTGT
CGATGGGGGA GGTCAAGAAG GATATAAAAA TAACAGTGAT TGCTTTTGTT ATCAATTATC
TGTTCTTTTA TATTCCGGTG TCATTATATC TTAGTTATTA CTATGGATAT AATTTTTTTA
ATCTATATAT GTTTTTTTTA TCACTTGTAG TTACATTTTT ATCGTTGTGG TTAAACGTGA
ATTTTACTT CTTACAAAT CTTATAGCGA AGGTGTTGAA ATGAGAAAC TATCTGAAAA
TGAAATAAAA CAAATATCTG GAGGTGACGG GAATGACGGG CAGGCAGAAT TAATTGCTAT
TGGTTCACCT GCTGGTACGT TTATTAGCCC GGGATTTGGT TCTATTGCAG GGGCTTATAT
AGGTGATAAA GTACATTCAT GGGCAACGAC TGCACGGGT AGTCCCTCCA TGTCTCCCTC
AGGTATAGGA TTATCATCCC AGTTTGGATC CGGCAGAGGT ACATCAAGTG CCTCTTCGTC
TGCGGGGAGT GGAAGTTAAA CCTTATATTG TTAATGAAGC ACCTGTTACA GGTGCTTCAT
TATATGATAA TGGTGCCTAT GTAACGCTTT AAGTTAGGCT CCTCCGAACG TTTATTATTT
TATGCATACC GCATAGTAAC AATGCCAGTC CGTTTAATTT TTATCATATG AGATTGTAAC
AGAAAATGCT CTGTTAATGA GTATAACTGA TAGAAGCCAG ATATTGTATT CCTGCTATGA
CTGCATTCAT CGACATGTGA AGTAATATGG GCATTAACAG GCCATTCGAC TTTATTTTTG
CACTAATAAG AACTAACGAT ACCAGAAAGA GCATCAAAAA TGTATGAAAG TCAGTATATT
GTAGATGTAA TGCAGAAAAT ATGACTGATG TTACAATAGC AGAGGTGTAA ATATTGTCAT
TAAACCAGAA TTTAAAAGAG TTAAAAGAC ACCCTCGAAA TACAATCTCT TCATATACAG
GAACAAGAAG TACAGAAGAA AATATATTAA TCCATAAAAT GGCTCCGGAA AAAGACATTT
GCGAGATCAT CCATTCTTCC GTTTTCCTCA CTGCCAGCAG ATAAGGGACA AATAACTGAA
TAATCATTAT TAAAGAGAAA AGAGTAAAAA AAACATCCAG CCGAAAAGAA CCTCTTCCTA
ATCCTTCTCT TTTTCGGAAG AAAATAAAAT ACAGTGGAAC AAGAATTAGA AATTCAGCAA
GAAATAATGC TGGAACAAGA AGTCCCCTGG ATATAAGTTC CTGCCTGTTT GTAAGAAATG
CAGGAATAAA GGTAATAGAA AATGATAAAG TAAACATTGA AAAGCAGCCG GCTGATGACT
GTTTTTTTAT AGTTGCATTT TTATTTTTTG TCATTTCCAT T

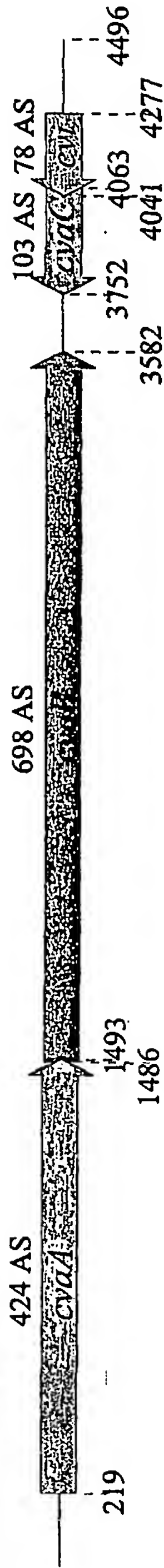
```

Abbildung 1a Organisation des Gen-Clusters der mcm-Region aus E. coli DSM 6601 mit den für die Microcin-Synthese, -Prozessierung und -Ausscheidung notwendigen Genabschnitten (unten) und Vergleich mit den Gen-Clustern für die Produktion von Microcin 24 (oben) und Colicin V (mitte)

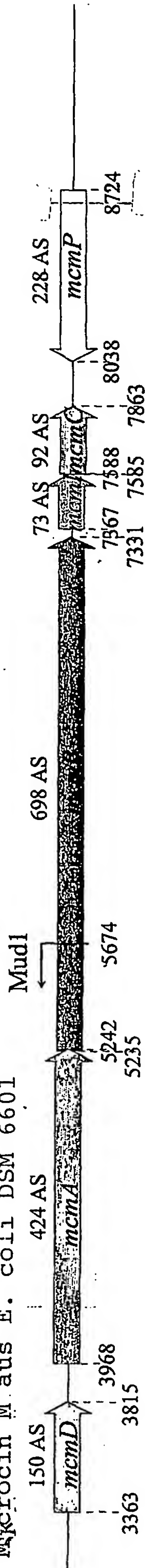
Microcin 24 Acc: U47048



Colicin V Acc: X57524/X57525



Microcin M aus E. coli DSM 6601



lacZ-Promotor pUC19 →

3967

→ Cval-Pr.

7990

← Cval3-Pr.

pSP112/22

Abbildung 2 Aminosäuresequenz des antimikrobiellen Peptids
(Microcin M, MmcC) aus *E. coli* DSM 6601

Protein 1	DGNDGQAELIAIGSLAGTFISPGFGSLAGAYIGDKVHS WATTATVSPSMSPSGIGLSSQFGSGRGTSSASSSAGSGS
-----------	---

Abbildung 3 Überlappende Oligopeptide (Teilsequenzen) aus
dem antimikrobiellen Peptid (Microcin M, MmcC)
aus *E. coli* DSM 6601

Peptide	Mol. Weight (mol/g)	Weight (mg)	mmol	Mixture	Protein
1. DGNDGQAELIAIGSLA	1543.7	0.46	0.0003	M.1	1
2. GQAELIAIGSLAGTFI	1560.8	0.47	0.0003	M.1	1
3. LAIGSLAITFISPGFG	1507.8	0.45	0.0003	M.1	1
4. LAGTFISPGFGSIAGA	1465.7	0.44	0.0003	M.1	1
5. ISPGFGSIAGAYIGDK	1552.8	0.47	0.0003	M.1	1
6. GSIAGAYIGDKVHSPA	1631.3	0.49	0.0003	M.1	1
7. AYIGDKVHSPAATTATV	1719.9	0.52	0.0003	M.1	1
8. KVHSPAATTATVSPSMS	1639.9	0.51	0.0003	M.1	1
9. ATTATVSPSMSPSGIG	1463.6	0.44	0.0003	M.1	1
10. VSPSMSPSGIGLSSQF	1580.8	0.47	0.0003	M.1	1
11. SPSSGIGLSSQFGSGRG	1493.6	0.45	0.0003	M.1	1
12. GLSSQFGSGRGTTSSAS	1485.6	0.45	0.0003	M.1	1
13. FGSGRGTTSSASSSAGS	1402.4	0.42	0.0003	M.1	1
14. SGRGTSSASSSAGSGS	1342.3	0.40	0.0003	M.1	1

SEQUENZPROTOKOLL

<110> Pharma-Zentrale GmbH

<120> DNA-Sequenz und Microcin aus Escherichia coli Stamm DSM
6601

<130> PT66/000

<140> DE 10055789.9

<141> 2000-11-10

<160> 1

<170> PatentIn Ver. 2.1

<210> 1

<211> 5381

<212> DNA

<213> Escherichia coli

<400> 1

```

ataagggaaa tagtaatcat gtcttatata agggaaacca tcagaggaaa agatgaatgg 60
actgttttatg aacagatcgg ttttgcggtc agttgtatgc tctacaatcg taattacagt 120
ctgtatccgg tgttaaccat tcaatactgg actgaatatg cgatacagca taatcagatt 180
aaattcctgt ttgattcacg aggttttcca ctggcggtata ttacctgggc atatccttgag 240
gctgatacag aagcacgcct gctcagggat ccagaattca gggtgcatcc gtctgaatgg 300
aatgaagatg gaaggatctg gattctggat ttctgttgta aaccaggctt tggtcgaaaa 360
gttattgact atctcataca gcttcagcca tggggagaag gagaagtacg atgggtaagc 420
aggcgaaaga aaattgtgac atacatccct gagcggctgc ataaaacgta gtacctctga 480
agatacagag ataattgtaa attacggggg aaatgcacgc ctgatactat tttgacagga 540
ctctgtatgt tctggctttg attatagcat ttcggttaac cagtgaagggt attacaggaa 600
aagtgtgagt aagaggagct attttggttc gtcaggatgc tttagaaaac agaaaaatga 660
agtggcaggg acgggcaata ttacttcccg gaataccact atgggttaatc atgctgggaa 720
gcatagtgtt tattacggca tttctgatgt tcattattgt tgggtacctat agccgccgtg 780
ttaatgtcag tgggtgaggtc acaacctggc caagagctgt caatatatat tcagggtgtac 840
agggatttgt tgtcaggcag tttgttcatg aaggacagtt gataaaaaaa ggggatcctg 900
tttatctgat tgacatcagt aaaagtacac gtagtggtat tgtcactgat aatcatcgtc 960
gggatataga aaaccagctg gttcgtgtgg acaagattat ttcccgctctg gaagaaagta 1020
aaaaaataac gctggatacc ctggaaaaac aacgcctgca atacacagat gcgtttcgtc 1080
gctcatcaga tattatacag cgtgcagagg aagggataaa aataatgaaa aacaacatgg 1140
agaattacag aaactatcag acaaaagggc tgattaataa agatcagtta actaaccagg 1200
tggcattata ttatcagcaa caaaacaatc ttcttagcct gagcggacag aacgaacaga 1260
atgccctgca gataaccact ctggagagtc agattcagac tcaggcagca gattttgata 1320
accgtatcta ccagatggaa ctgcaacggt acgagttaca gaaagaactg gttaacactg 1380
atgtggaggg cgaaatcatt atccgggctg tgactgacgg gaaagttgac tccctgagtg 1440
tactgtcgg gcaaattggtc aatcccggag acaaccttct gcaggttatt cctgagaaca 1500
ttgaaaatta ttatcttatt ctctgggtcc cgaatgatgc tgttccttat atttcggctg 1560
gtgacaaagt gaatattcgt tatgaagcct ttccggcaga aaaatttggg cagttttctg 1620
ctacgggttaa aactatatcc aggactcctg catcaacaca ggaaatgttg acctataagg 1680
gagcaccaca gaatacgccg ggcgcctctg ttccctggta taaagtcatt gcgatgcctg 1740

```

aaaagcagat	tatcagatat	gacgaaaaaa	acctccctct	ggaaaatgga	atgaaagccg	1800
aaagtacact	atcttctggaa	aaaaggcgta	tttaccagtg	gatgctttct	cctttctatg	1860
acatgaaaca	cagtgcaca	ggaccgctca	atgactaacg	ggaatttcag	acaaattata	1920
aatcagcttg	atatgcgctg	gcgacgtcgt	gttccgggtta	ttcatcagac	ggagaccgct	1980
gaatgtggac	tggcctgcct	ggcaatgata	tgcgggtcatt	ttggtaagaa	tattgacctg	2040
atatctcttc	gtcgtaagtt	taatctctcg	gcccggtggag	caaacccttg	aggaatcaat	2100
ggaatagcgg	agcagctggg	gatgatcacc	cgggctcttt	cactggagct	ggatgaactt	2160
ggtgccctca	aaatgccgtg	tattctccac	tgggatttca	gtcacttcgt	cgtgctggtc	2220
agcgtaaagc	gtaaccgtta	tgtactgcat	gatccggcca	gaggaagaag	atatctcggt	2280
cgggaggaaa	tgagccggta	ttttacgggc	attgcacttg	aggtctggcc	cggagggtgaa	2340
ttccggacgg	aaacccagca	aacccgcata	agtctccggt	cactgattaa	cagtatttac	2400
ggtattaaaa	gaacactggc	gaaaattttc	tgtctgtcag	ctgtaattga	agcaatcaat	2460
ctggtaatgc	cgggtgggaac	tcagctgggt	atggatcatg	cgattccggc	gggggaccga	2520
gggctgctga	cgcttatttc	tgctggcctg	atgttcttta	tattgctcag	ggccgcgggtg	2580
agtatgctgc	gtgcatggtc	ctcactgggt	atgagcacgc	tcataaatat	acagtggcag	2640
tcgggtctgt	ttaaccacct	tctcagactg	ccgctggcat	tttttgaacg	ccgtaaatta	2700
ggtgatatcc	agtcgcggtt	tggctccctt	gacactttga	gggccacctt	taccacctgt	2760
gtggttgggg	caatcatgga	cagtattatg	gttgtgggtg	tttttgtgat	gatgctgtta	2820
tatggaggat	atcttacctg	gatagttctc	ggttttacca	tggtttacgt	tcttattcgt	2880
ctggtgacat	acggctatta	ccggcaaata	tcggaagaaa	ctcttgtcag	gggggcccgg	2940
gccagctcct	atcttatgga	aagcctgtat	ggtattgcca	cggtaaaaat	ccagggtatg	3000
gtcgggatcc	ggggaacaca	ctggcttaac	ctgaaaatag	atgcgatcaa	ttcagggtatt	3060
aagttaacca	ggatggattt	gctcttcggg	ggaataaata	cttttgttgc	cgctgtgat	3120
cagggtggcga	ttttatggct	gggtgcaagc	cttgtgatcg	ataatcagat	gacaataggg	3180
atgtttgtgg	catttggttc	ttttcgtggg	cagttttcgg	atcggggttg	ttcgctgacc	3240
agttttcttc	ttcaactgag	aataatgagt	ctgcataatg	agcgcattgc	agataattgca	3300
ctacatgaaa	aggaagaaaa	gaaaccggaa	attgaaatcg	ttgctgacat	gagcccgggt	3360
tcactggaaa	ccactgattt	aagctaccgg	tatgacagcc	agtcagcaca	ggtattcagt	3420
ggtctgaatt	tgtctgtggc	tccgggagaa	agtgtggcta	taactgggtg	ctccgggtgcc	3480
ggaaaaacca	cattaatgaa	agtattatgt	ggactgtttg	aaccagatag	tggaaaagta	3540
ctggttaatg	gcacggatat	acgtcaactt	ggaataaata	attatcaccg	tatgatagcc	3600
tgtgttatgc	aggacgaccg	gctattttca	ggatcaattc	gtgaaaatat	ctgtgggttt	3660
gcagaagaaa	cagacgacga	atggatgaca	gaatgtgcc	gagcaagtca	tattcatgat	3720
gtgataatga	aaatgccaat	ggggtatgaa	acgttaatag	gtgaactggg	ggaagggtctt	3780
tccggcggtc	aaaaacagcg	tatattcatt	gcccgagctt	tataccggaa	acctggaata	3840
ttatttatgg	atgaggctac	aagttctctt	gatacagaaa	gtgaacgttt	cgttaatgct	3900
gccataaaaa	aaatgaatat	caccggggtg	attattgcac	acagagaaac	tacgttgaga	3960
actgttgaca	ggattatttc	tattttaa	ccactgggtg	aactttgtaa	ggagttttgt	4020
cgatggggga	ggtcaagaag	gatataaaaa	taacagtgat	tgcttttgtt	atcaattatc	4080
tgttctttta	tattccgggtg	tcattatatc	ttagtatta	ctatggatat	aattttttta	4140
atctatatat	gtttttttta	tcacttgtag	ttacattttt	atcgttgttg	ttaaacgtga	4200
atttttactt	cttcacaaat	cttatagcga	aggtgttgaa	atgagaaaac	tatctgaaaa	4260
tgaaataaaa	caaatatctg	gaggtgacgg	gaatgacggg	caggcagaat	taattgctat	4320
tggttcactt	gctgggtacgt	ttattagccc	gggatttggt	tctattgcag	gggcttatat	4380
aggtgataaa	gtacattcat	gggcaacgac	tgcgacgggt	agtcctcca	tgtctccctc	4440
aggtatagga	ttatcatccc	agtttggatc	cggcagaggt	acatcaagtg	cctcttcgtc	4500
tgcggggagt	ggaagttaaa	ccttatattg	ttaatgaagc	acctgttaca	ggtgcttcat	4560
tatatgataa	tgggtgcctat	gtaacgcttt	aagttaggct	cctccgaacg	tttattattt	4620
tatgcatacc	gcatagtaac	aatgccagtc	cgtttaattt	ttatcatatg	agattgtaac	4680
agaaaatgct	ctgttaatga	gtataactga	tagaagccag	atattgtatt	cctgctatga	4740
ctgcattcat	cgacatgtga	agtaatatgg	gcattaacag	gccattcgac	tttatttttg	4800

cactaataag	aactaacgat	accagaaaga	gcatcaaaaa	tgtatgaaag	tcagtatatatt	4860
gtagatgtaa	tgcagaaaaat	atgactgatg	ttacaatagc	agaggtgtaa	atattgtcat	4920
taaaccagaa	tttaaaagag	ttaaaaagac	accctcgaaa	tacaatctct	tcatatacag	4980
gaacaagaag	tacagaagaa	aatatattaa	tccataaaaat	ggctccggaa	aaagacattt	5040
gcgagatcat	ccattcttcc	gttttcctca	ctgccagcag	ataagggaca	aataactgaa	5100
taatcattat	taaagagaaa	agagtaaaaa	aaacatccag	ccgaaaagaa	cctcttccta	5160
atccttctct	ttttcggaag	aaaataaaaat	acagtggaac	aagaattaga	aattcagcaa	5220
gaaataatgc	tggaacaaga	agtcccctgg	atataagttc	ctgcctgttt	gtaagaaatg	5280
caggaataaaa	ggtaatagaa	aatgataaag	taaacattga	aaagcagccg	gctgatgact	5340
gttttttttat	agttgcattt	ttattttttg	tcatttccat	t		5381